



## **RAW SEQUENCE LISTING** **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/002,536  
Source: OPE  
Date Processed by STIC: 8/2/02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

# Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 10/002,536
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 _____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 _____ Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 _____ Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 _____ Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 _____ Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 _____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 _____ Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 _____ Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 _____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 _____ Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 <input checked="" type="checkbox"/> Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 _____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 _____ Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.	



**Does Not Comply**  
**Corrected Sequence Needed**

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/002,536

DATE: 08/02/2002  
TIME: 16:24:53

Input Set : A:\PTO.VSK.txt  
Output Set: N:\CRF3\08022002\J002536.raw

3 <110> APPLICANT: Kane, Michael D.  
4 Nagel, Aaron C.  
5 Dombkowski, Alan A.  
7 <120> TITLE OF INVENTION: COMPOSITIONS AND SYSTEMS FOR IDENTIFYING AND COMPARING  
EXPRESSED GENES  
8 (mRNAs) IN EUKARYOTIC ORGANISMS  
10 <130> FILE REFERENCE: 65446-87  
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/002,536  
C--> 12 <141> CURRENT FILING DATE: 2001-11-01  
12 <160> NUMBER OF SEQ ID NOS: 5  
14 <170> SOFTWARE: PatentIn version 3.1  
16 <210> SEQ ID NO: 1  
17 <211> LENGTH: 61  
18 <212> TYPE: DNA  
19 <213> ORGANISM: Artificial Sequence  
21 <220> FEATURE:  
22 <223> OTHER INFORMATION: This is an artificial sequence.  
24 <220> FEATURE:  
25 <221> NAME/KEY: misc\_feature  
26 <222> LOCATION: (1)..(50)  
27 <223> OTHER INFORMATION: The poly-T tail may have between 8 and 50 residues.  
30 <220> FEATURE:  
31 <221> NAME/KEY: misc\_feature  
32 <222> LOCATION: (51)..(51)  
33 <223> OTHER INFORMATION: v may be selected from a or c or g.  
36 <220> FEATURE:  
37 <221> NAME/KEY: misc\_feature  
38 <222> LOCATION: (52)..(62)  
39 <223> OTHER INFORMATION: The sequence may have between 3 and 10 residues selected from  
a o  
40 r c or g or t.  
43 <400> SEQUENCE: 1  
W--> 44 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt vnnnnnnnnnn 60  
W--> 46 n 61  
49 <210> SEQ ID NO: 2  
50 <211> LENGTH: 25  
51 <212> TYPE: DNA  
52 <213> ORGANISM: Artificial Sequence  
54 <220> FEATURE:  
55 <223> OTHER INFORMATION: This sequence is artificial.  
57 <220> FEATURE:  
58 <221> NAME/KEY: misc\_feature  
59 <222> LOCATION: (23)..(25)  
60 <223> OTHER INFORMATION: n may be selected from a or c or g or t.

The type of errors shown exist throughout  
the Sequence Listing. Please check subsequent  
sequences for similar errors.

→ must explain genetic source, see  
error summary sheet item 11

63 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 08/02/2002

PATENT APPLICATION: US/10/002,536

TIME: 16:24:53

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\08022002\J002536.raw

```

64 <221> NAME/KEY: misc_feature
65 <222> LOCATION: (24)..(24)
66 <223> OTHER INFORMATION: v may be selected from a or c or g.
69 <400> SEQUENCE: 2
W--> 70 tttttttttt tttttttttt tvnnnn 25
73 <210> SEQ ID NO: 3
74 <211> LENGTH: 24
75 <212> TYPE: DNA
76 <213> ORGANISM: Artificial Sequence
78 <220> FEATURE:
79 <223> OTHER INFORMATION: This is an artificial sequence.
81 <400> SEQUENCE: 3
82 tttttttttt tttttttttt aaac 24
85 <210> SEQ ID NO: 4
86 <211> LENGTH: 24
87 <212> TYPE: DNA
88 <213> ORGANISM: Artificial Sequence
90 <220> FEATURE:
91 <223> OTHER INFORMATION: This is an artificial sequence.
93 <400> SEQUENCE: 4
94 gtttaaaaaa aaaaaaaaaa aaaa 24
97 <210> SEQ ID NO: 5
98 <211> LENGTH: 25
99 <212> TYPE: DNA
100 <213> ORGANISM: Artificial Sequence
102 <220> FEATURE:
103 <223> OTHER INFORMATION: This is an artificial sequence.
105 <400> SEQUENCE: 5
106 tttttttttt tttttttttt ggttt 25

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/002,536

DATE: 08/02/2002  
TIME: 16:24:55

Input Set : A:\PTO.VSK.txt  
Output Set: N:\CRF3\08022002\J002536.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 52,53,54,55,56,57,58,59,60,61  
Seq#:2; N Pos. 23,24,25

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/002,536

DATE: 08/02/2002

TIME: 16:24:55

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF3\08022002\J002536.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No  
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:44 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0  
L:46 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:60  
L:70 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0